Modeling

Motivation

With the rapid development of synthetic biology, there is an increasing requirement of an accurate quantitative regulation of gene expression. Moreover, researchers can provide insight into how the quantitative regulation system can be improved with the application of mathematical modeling.

Briefly, our team has developed a brand new method of accurate regulation on post-transcriptional level by means of utilizing the effect of inhibiting degradation of mRNA stem loop. To validate the effectiveness of initial designs, we built a mathematical modeling to simulate the designed system and mutually authenticate the theoretic data with the experimental data with the purpose of not only guiding both our design and experiment but also selecting the appropriate parameters of modeling to produce the most ideal theoretical data.

Introduction

Fundamentally, our gene circuit design of stem loops share a gradient free energy. So firstly, we seek for Michaelis-Menten kinetics, partition function, empirical formula (literature material) to respectively determine the constant and parameter of our model. Secondly, in order to simulate the whole system, we expect to determine the mathematical expression between the values of free energy of different stem loops with the constant of the corresponding constant of decay. To solve this problem, we seek for the QSAR/QSPR method (quantitative structure-activity relationship/quantitative structure–property relationships) combined with experimental data. Finally, we simulate the system and expect to improve our design to make it realistic and practical.

Metabolic reaction networks

$$\emptyset \xrightarrow{K_r} [\mathrm{mRNA}] \tag{1}$$

$$[\mathrm{mRNA}] \xrightarrow{\mathrm{K}_{p_1}} [\mathrm{mRNA}] + [\mathrm{Protein}_1] + [\mathrm{Protein}_2] \tag{2}$$

$$[\mathrm{mRNA}] \xrightarrow{\mathrm{K}_{\mathrm{d}_1}} [\mathrm{mRNA}_1] + [\mathrm{mRNA}_2] \tag{3}$$

$$[\mathrm{mRNA}_1] \xrightarrow{\mathrm{K}_{\mathrm{p}_{11}}} [\mathrm{mRNA}_1] + [\mathrm{Protein}_1] \tag{4}$$

$$[\mathrm{mRNA}_2] \xrightarrow{\mathrm{K}_{p_{12}}} [\mathrm{mRNA}_2] + [\mathrm{Protein}_2] \tag{5}$$

$$[\mathrm{mRNA}] \xrightarrow{\mathrm{K}_{d_0}} \emptyset \tag{6}$$

$$[\mathrm{mRNA}_1] \xrightarrow{\mathrm{K}_{d_{11}}} \emptyset \tag{7}$$

$$[\mathrm{mRNA}_2] \xrightarrow{\mathrm{K}_{\mathrm{d}_{12}}} \emptyset \tag{8}$$

$$[\mathrm{Protein}_1] \xrightarrow{\mathrm{K}_{\mathrm{d}_{p_1}}} \emptyset \tag{9}$$

$$[\mathrm{Protein}_2] \xrightarrow{\mathrm{K}_{\mathrm{d}_{p_2}}} \emptyset \tag{10}$$

symbol description

Symbol	Definition	Units
$[\mathrm{mRNA}]$		
$[\mathrm{mRNA}_1]$		

$[\mathrm{mRNA}_2]$		
$[\mathrm{Protein}_1]$		
$[\mathrm{Protein}_2]$		
K_r		
K_{p_1}		
$\mathrm{K}_{\mathrm{d}_1}$		
$\mathrm{K}_{\mathrm{p}_{11}}$		
$\mathrm{K}_{p_{12}}$		
K_{d_0}		
$\mathrm{K}_{d_{11}}$		
$\mathrm{K}_{\mathrm{d}_{12}}$		
$\mathrm{K}_{\mathrm{d}_{p_1}}$		
$\mathrm{K}_{\mathrm{d}_{p_2}}$		

Assumptions and hypothesis

Model development

1) Michaelis-Menten kinetics

To determine the reaction constant $\mathrm{K_r}$ in (1), we seek for Michaelis-Menten kinetics. Further, we take the pilot process of DNA transcription into consideration for the sake of validating the accuracy of calculation and finally got satisfactory results.

symbol	definition	units
[AraC]	The concentration of dissociative repressor protein - AraC	
[Arab]	The concentration of arabinose	
$[AraC \cdot Arab]$	The concentration of complex - $AraC \cdot Arab$	
$[AraC]_T$	The sum of the concentration of both dissociative repressor protein-AraC and complex - $AraC \cdot Arab$	
$k_i, i = 1,2,3$	reaction rate constant	
k_m	Michaelis constant	
v	transcription rate	

In our circuit design, we chose araBAD promoter, which will be combined with repressor protein - AraC and the latter represses transcription of mRNA without arabinose. Then, Arabinose of reagent addition will bind to AraC and form the Arab • AraC compound, allowing transcription to occur.

【hypothesis】We make an assumption that AraC is always in large concentration and that its binding to arabinose happens on a faster time scale to transcription. Therefore, we do not need to consider the individual concentrations of arabinose and AraC, instead we just need to include the concentration of the complex Arab • AraC.

The process boils down to following formula:

$$\mathrm{AraC} + \mathrm{Arab} \rightleftharpoons_{k_2}^{k_1} \mathrm{AracC} \cdot \mathrm{Arab} \xrightarrow{k_3} \mathrm{mRNA} + \mathrm{AraC}$$

according to law of mass action,

$$\frac{d[AraC \cdot Arab]}{dt} = k_1 \cdot ([AraC]_T - [AraC \cdot Arab]) \cdot [Arab] - k_2 \cdot [AraC \cdot Arab] - k_3 \cdot [AraC \cdot Arab]$$

$$\because \frac{d[AraC \cdot Arab]}{dt} = 0 \quad \therefore \frac{k_2 + k_3}{k_1} = \frac{([AraC]_T - [AraC \cdot Arab]) \cdot [Arab]}{[AraC \cdot Arab]}$$

Define $k_m = \frac{k_2 + k_3}{k_1}$,then

$$[AraC \cdot Arab] = \frac{[AraC]_T \cdot [Arab]}{k_m + [Arab]}$$

Then the transcription rate can be confirmed like following mathematical expression:

$$v = k_3 \cdot [AraC \cdot Arab] = k_3 \cdot [AraC]_T \cdot \frac{[Arab]}{k_m + [Arab]}$$

Further, according to the theory of order of reaction, transcription rate can be convert into reaction rate constant.

$$\mathrm{K_r} = \frac{\mathrm{v}}{[AraC \cdot Arab]}$$

2) Partition function

To determine the reaction rate constant $\mathrm{K_{p_1}}, \mathrm{K_{p_{11}}}, \mathrm{K_{p_{12}}}$, we seek for partition function. Inspired by the references [ ], we applied the partition function to the dynamic description of the translation process, in which we can obtain the probability that the ribosomes bind to RNA and then fortunately succeeded in converting the probability into translation rate. Further, we can calculate the reaction rate constant $\mathrm{K_{p_1}}, \mathrm{K_{p_{11}}}, \mathrm{K_{p_{12}}}$ by utilizing the concentration data of (2),(4),(5) from our wet lab experiments.

【hypothesis】Number of ribosomes is always in large concentration and ribosomes are more inclined to bind with RBS.

symbol	definition	units
P_{bound}	Probability of ribosome binding to RBS	/
P	Effective number of ribosome available for binding to RBS	
N_{NS}	The number of nonspecific site of mRNA	
K_{pd}^{S}	Dissociation constants for specific binding	nM
K_{pd}^{NS}	Dissociation constants for non-specific binding	nM
ε_{pd}^{S}	Binding energy for ribosome on the RBS	J
ε_{pd}^{NS}	Average binding energy of ribosome to the genomic background	J
k_B	Boltzmann constants	/

T	Temperature	K
Rate	Rate of reaction	
Volume	Volume	L
Avogadro	Avogadro constants	/
[mRNA](0)	Initial concentration of mRNA	

According to the document literature [], the probability that the ribosomes bind to RNA can be obtained as following:

$$P_{bound} = \frac{1}{1+\frac{N_{NS}}{P}\exp(\frac{\varepsilon_{pd}^{S}-\varepsilon_{pd}^{NS}}{k_B T})}$$

Because $\varepsilon_{pd}^{S}-\varepsilon_{pd}^{NS} \approx k_B T \ln(\frac{K_{pd}^{S}}{K_{pd}^{NS}})$, therefore the expression above can be further simplified into the following:

$$P_{bound} = \frac{1}{1+\frac{N_{NS}}{P}\cdot\frac{K_{pd}^{S}}{K_{pd}^{NS}}}$$

Then, convert the probability into reaction rate [ ]:

$$Rate = \frac{1000\times P_{bound}}{Volume\times Avogadro}$$

Then, to integrate experimental data and emulation analysis, we use the method of curve fitting to mRNA concentration in the initial time. Finally, we obtain the rate constant:

$$k = \frac{1000\times P_{bound}}{Volume\times Avagadro\times [mRNA](0)}$$

3) Empirical formula

To determine the reaction rate constant of mRNA cleavage by RNase E, we chose the empirical formula, obtained from document literature.

$$K_{d_1} = \frac{[H^+]K_{E1}k_0}{K_{E1}\cdot K_{E2}+[H^+]K_{E1}+[H^+]^2}$$

4) QSAR/QSPR

One of our goals is to determine the mathematical expression between the value of free energy of different stem loops with $K_{d_0}, K_{d_{11}}, K_{d_{12}}$ —— the constant of the corresponding constant of decay. we seek for the QSAR/QSPR method (quantitative structure-activity relationship/quantitative structure–property relationships) combined with experimental data.

Principal steps of QSAR/QSPR including

(i) Selection of Data set and extraction of structural/empirical descriptors

(ii) variable selection

(iii) model construction

(iv) validation evaluation.

Hansch-equation $\lg\frac{1}{C} = \lg A - 0.434\frac{\Delta G}{RT}$

symbol	definition
C	
A	
ΔG	
R	
T	

$$\lg\frac{1}{C} = -a\pi^2 + b\pi + \rho\sigma + \delta E_S + c$$

a) Hydrophobic parameter

$$\lg P = \lg P_H + \sum(\pi x_i)$$

b) Electronic effect parameter

$$\lg\frac{k_X}{k_H} = \rho\sigma_X$$

c) Steric parameter

$$MR = \frac{(n^2 - 1)M_W}{(n^2 + 2)d}$$

d) Other parameter

Evaluation system of QSAR model

$$r = \sqrt{1 - \frac{\sum(Y_{cal} - Y_{exp})^2}{\sum(Y_{exp} - \overline{Y}_{exp})^2}}$$

$$s = \sqrt{\frac{\sum(Y_{cal} - Y_{exp})^2}{n - k - 1}}$$

$$F = \sqrt{\frac{r^2(n - k - 1)}{k(1 - r)^2}}$$

Simulation

At the beginning of the simulation part, let's replace the parameters with simple characters.

We chose the Gillespie algorithm to get the Exact Stochastic Simulation of Coupled.

Now, we are ready to simulate the system by Matlab.

tspan: Initial and final times, [t_init, t_final].

x0: Initial species amounts, [S1_0, S2_0, ...].

stoich_matrix: Matrix of stoichiometries (Nreactions x Nspecies).

prop_fcn: Function that calculates reaction propensities.

params: User-defined parameters, passed to prop_fun

Conclusion

References